

Disruption of overlearned discriminative behavior in monkeys (*Cebus apella*) by delay of reward

JACQUILYN K. COX and M. R. D'AMATO

Rutgers University, Busch Campus, Rutgers University, New Brunswick, New Jersey 08903

Previous work has shown that when a delay of reward (DOR) is introduced into a well-learned discrimination, even gradually, discriminative performance deteriorates and, with moderately long DORs, does not recover with practice. The present experiment assessed whether the decrement in performance was due to an associative loss or to a decline in the incentive value of the reward object caused by the DOR. Cebus monkeys were trained on a simple visual discrimination and tested with either a DOR or an identical delay period which preceded the appearance of S+ and S- ("predelay" trials); reinforcement on predelay trials was immediate. On half of the daily trials, the animals were given the option of choosing either the DOR or the predelay trial. The duration of the delay was increased gradually until terminal delays of 32 to 128 sec were reached. All four animals maintained almost errorless performance on predelay trials; in contrast, their error rate reached 36% on DOR trials. Surprisingly, none of the animals learned to choose predelay over DOR trials. Both results were interpreted in terms of the incentive loss hypothesis.

Delay of reward (DOR) has been a variable of interest to learning psychologists for more than half a century (cf. Mowrer, 1960, Chap. 10; Renner, 1964). Much of the research effort during this time has been devoted to assessing the effects of DOR on the acquisition of behavior, and until recent years its influence on the performance of well-established instrumental behaviors went largely unnoticed. Research conducted during the last decade has shown that introducing a DOR into an already established simple instrumental response usually causes a serious disruption of performance, which is often long lasting (e.g., Bullock & Richards, 1973; Ferster & Hammer, 1965).

In the case of simple instrumental behaviors, the deficit caused by DOR relates to response vigor—reductions in response rate and speed, and increases in latency. It has recently been shown, however, that a DOR, even when introduced gradually, can cause a lasting deterioration in response accuracy. In a lengthy series of experiments, D'Amato and Cox (1976) found that a DOR introduced gradually into a well-learned visual discrimination caused a loss of response accuracy, sometimes to the point where the animals responded more to S- than to S+. Even after many thousands of test trials distributed over almost 2 years, their monkeys were unable to maintain a high performance level on a heavily overlearned discrimination with DORs in the neighborhood of a minute or so. These results, which very

likely are not restricted to cebus monkeys (cf. Mishkin & Weiskrantz, 1958), have a direct bearing on the role of DOR in the acquisition of discriminative behaviors. If animals cannot maintain a well-learned visual discrimination under the burden of a moderate DOR, it seems unlikely that they will be able to acquire new discriminative behaviors when so disadvantaged.

The present experiment was addressed to two issues, one empirical and one theoretical. Animals were exposed to immediate and delayed reward trials on the same, well-learned, visual discrimination. On immediate reward trials, an identical delay occurred as on DOR trials, but it preceded the presentation of S+ and S-. The two types of trials (DOR and "predelay") were randomly intermixed, although the animal could discriminate at the time of choice which condition was in effect. With regard to the empirical question, we wished to determine whether DOR would result in the usual decline in discriminative performance even though the animals maintained contact with immediate reinforcement on the same discriminative task.

As for the theoretical issue, the results of the experiment should have a bearing on two plausible explanations of why DOR causes disruption of well-established discriminative behaviors. One interpretation is that when a discriminative response is separated from its consequences by a DOR and no differential cues exist to mediate the delay interval, the association between the two events becomes weakened, a position that shall be referred to as the "associative loss" hypothesis. The second interpretation is that the performance decrement is due to a loss of incentive value caused by the DOR. Accord-

ing to this view, the incentive value of a reward object is inversely related to the delay that intervenes before its delivery (cf. Ainslie, 1975; Davenport, 1962; Logan, 1965). Moreover, a DOR can elicit negative emotional reactions, such as frustration, which may further depress incentive value (D'Amato & Cox, 1976). Consequently, an animal commits errors on a well-learned discriminative task after a DOR has been introduced, not because it has forgotten the contingencies between its responses and the subsequent outcomes (the associative loss hypothesis), but because the consequence of responding to S+ is insufficiently positive.

To relate the present experiment to these theoretical positions, we may assume from the associative loss hypothesis that the association between response and outcome is strengthened on predelay trials (because of immediate reinforcement) and weakened on DOR trials. Because of the intermixing of the two types of trials, the strength of association between response and outcome will depend upon the sequence of previous trials but will be the same, on average, for predelay and DOR trials. Consequently, discriminative performance should be comparable on both.

On the other hand, at the time that S+ and S- appear, the animal can discriminate the nature of the trial in progress and therefore whether reinforcement is immediate or delayed. This should be sufficient, from the incentive loss hypothesis, to generate a substantial difference in discriminative performance on predelay and DOR trials, which afford high and low incentives, respectively.

As a means of obtaining additional evidence bearing on the two hypotheses, the animals were frequently allowed to choose between predelay and DOR trials. In the event that the animals performed at a substantially higher level of accuracy on the predelay trials, the incentive loss hypothesis makes the interesting prediction that they will show little, if any, preference for predelay over DOR trials. The basis for this prediction is that the response through which the animal indicates its preference is separated from the terminal reinforcing events (reward pellet or time-out) of each trial by a period of time that is at least as long as the prevailing DOR. If the latter is such as to cause a substantial decline in performance on the visual discrimination, we may infer that it has reduced the incentive value of the reward object to a low level. If this is true for the (visual) discriminative response, it must also be true for the preference response, which is further removed in time from the terminal reinforcing events. Consequently, despite a higher density of reinforcement, there should be little incentive for the animals to acquire and maintain a preference for the predelay trials. On the other hand, the associative loss hypothesis, and common sense as well, leads to the

expectation that the animals will develop a preference for the more frequently rewarded alternative.

Motivated by a different theoretical issue, D'Amato and Cox (1976) reported a study that followed the design just described. Three of the four monkeys tested committed significantly more errors on DOR than on predelay trials when the delay reached 60 sec, the longest delay investigated. However, because only three or four sessions were run at the 60-sec delay, these results are only suggestive; moreover, because of the small number of sessions employed, no valid measure of preference for the two types of trials could be obtained.

METHOD

Subjects

Two adult female (Peanuts and Bridgette) and male (Hubert and Phurp) cebus monkeys (*Cebus apella*) served as subjects. All of the subjects except Bridgette had considerable past experience with DOR in a simple two-choice discrimination task. Throughout the experiment, the animals were maintained at 85%-93% of their full-ration body weight by appropriate daily rations of Purina Monkey Chow delivered 3-4 h following a test session. They were housed individually with free access to water in their home cages.

Apparatus

An on-line computer system (PDP 8/e, Digital Equipment Corp.) controlled stimulus selection, response contingencies, and data recording in all phases of testing and training. The computer was interfaced with two environmental primate chambers (Lehigh Valley Electronics), in which the animals were tested. One wall of the chamber was fitted with five Inline stimulus projectors (Industrial Electronics Engineers, Inc.). One projector was situated at each corner of a 12-cm square and the fifth at the center. Each projector was faced with a plastic key which served as the response mechanism. A microswitch, with which the animals could initiate a trial, was positioned 6 cm below the center projector. A response lever and a display unit, which was used to signal that the lever was operative, were located on each side of the projector array, separated by a distance of 4 cm.

Procedure

All subjects were trained on a simple two-choice simultaneous discrimination in which S+ was an inverted triangle and S- a vertical line. The subject initiated a trial by completing 15 responses on the microswitch, which produced S+ and S- on two of the four outer projectors (the center projector was not used). A correct response resulted in the immediate delivery of one 190-mg Noyes banana pellet, while an incorrect response produced a 1-min time-out signaled by a dimming of the overhead houselight. A response was ignored if it did not fulfill a 0.3-sec hold-down requirement. A 20-sec ITI, signaled by dimming of the overhead houselight and illumination of a light located below the projector array, followed delivery of the reward pellet or completion of the time-out. The animal was free to initiate the next trial at the termination of the ITI, which was indicated by full illumination of the houselight. Acquisition training continued until the criterion of 2 consecutive days with less than two errors per day was met, which required three to five 24-trial sessions, distributed one session per day. A 2-sec DOR was then introduced and an additional 6 to 14 24-trial sessions were given, at the end of which all subjects were responding at a high level of accuracy.

DOR vs. predelay performance. At this point the experiment proper began. Completion of the FR 15 on the microswitch now resulted in the illumination of the two display units above the response levers. One lever was assigned to deliver a standard

DOR-type trial; the other lever delivered a pre-delay trial. A response on the first lever resulted in the immediate appearance of S+ and S- on two of the projectors. A response to either stimulus produced the DOR, which was spent in total darkness. At the termination of the DOR, a pellet was delivered if the choice response had been correct; the usual 1-min time-out followed incorrect responses. A response on the second, pre-delay, lever immediately produced a delay interval of the same duration and stimulus conditions as the DOR, followed by presentation of S+ and S-. A response to S+ or S- led immediately to a reward pellet or a 1-min time-out, respectively. Thus the only difference between DOR and pre-delay trials was the positioning of the delay interval, which either preceded the response-outcome sequence (on pre-delay trials) or was interposed between response and outcome. It should be noted that, with regard to the lever response, on both types of trials response and outcome were separated by equivalent time intervals, which functionally served as DORs.

All animals were exposed to five delays in increasing order: 2, 4, 8, 16, and 32 sec. One animal (Peanuts) was taken to 64-sec and another (Phurp) to 64- and 128-sec delays. There were 16 24-trial sessions at each delay, and the position of the lever that delivered DOR trials was alternated from left to right every four sessions. To insure that the animals were exposed to both delay contingencies, on 6 of the daily 24 trials only the lever leading to DOR trials was operative (and its display unit illuminated) and on 6 other trials only the lever leading to the pre-delay trials was illuminated and operative.

Performance at terminal delays. The final phase consisted of exposing three of the animals to 20 24-trial sessions, during which the location of the lever delivering DOR trials was fixed on the side preferred by that animal. The delay interval was the longest interval reached in the previous phase, 32 sec for two of the animals and 64 sec for the third. Because of inordinately long response latencies on the levers and on the projectors, Phurp was dropped from this phase.

RESULTS

Figures 1 and 2 present the discriminative and leverpress behavior for each animal until the terminal delay was reached. Each data point is based on 16 test sessions. Examining first response accuracy on the visual discrimination, it is apparent that in all cases very few errors were committed on pre-delay trials, even at the longest delay. A high level of accuracy was also maintained on DOR trials at the short delay intervals. However, as the delay increased, accuracy on DOR trials slowly decreased, until a substantial difference emerged at the terminal delay. As evaluated by an analysis of variance performed on the four four-session blocks at the terminal delays, this difference was highly significant [$F(1/96) = 240.6, p < .001$]; moreover, individual t tests showed the difference to be significant for all animals ($p < .01$ in all cases).

On half of each session's 24 trials, the animal chose between pre-delay and DOR trials by pressing the appropriate lever. Because the level of accuracy on the two types of trials was not very different until the final delay was reached, there is little reason to expect the emergence of a preference for DOR or pre-delay trials prior to that point. The data presented in Figures 1 and 2 (open squares) confirm this

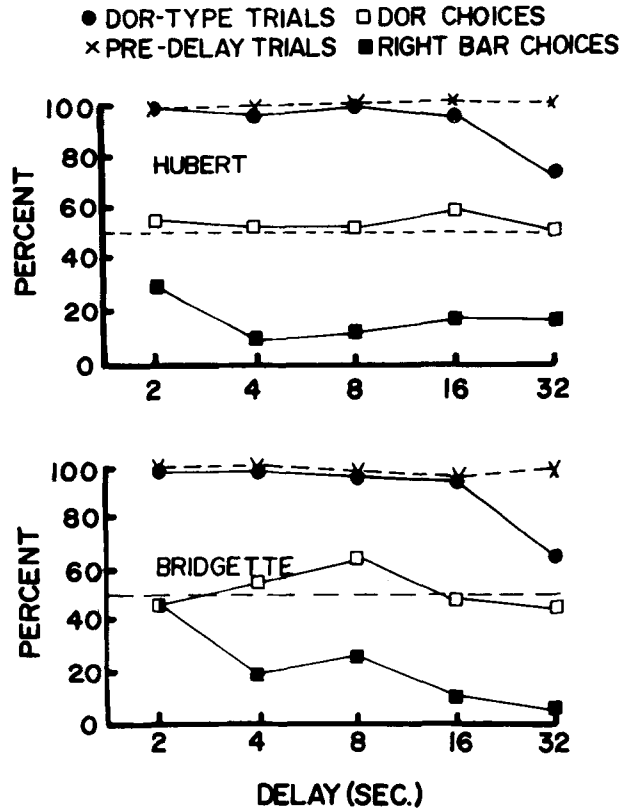


Figure 1. Performance on the visual discrimination and the lever (bar) response as a function of the delay interval and type of delay trial (pre-delay or DOR). Each point is based on 16 24-trial sessions, during which the lever delivering DOR-type trials was alternated from left to right every four sessions. The DOR choices (open squares) indicate the percentage of trials on which the animal chose the lever that led to a DOR-type trial.

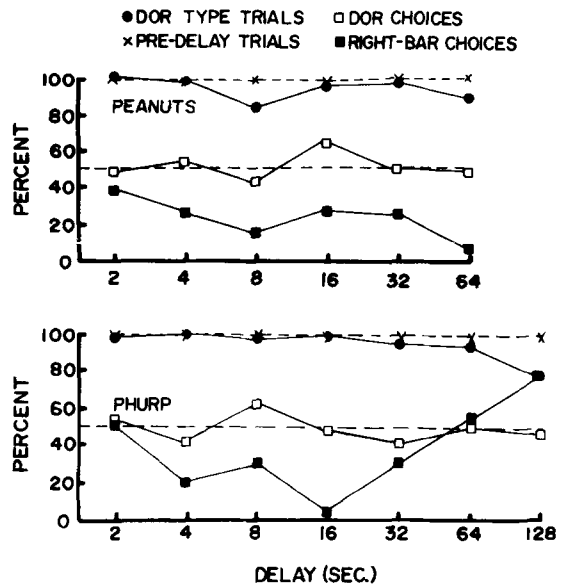


Figure 2. The visual discrimination and leverpress performance of the two animals that went beyond the 32-sec delay.

expectation. However, they also reveal that in all cases a preference failed to emerge during the 16 sessions at the terminal delay; the average percentage choice of the lever leading to DOR trials was very close to chance, 47.7%.

Three of the four animals developed a significant position preference, in all cases for the left lever. The fourth animal, Phurp, developed a similar preference, but midway through testing this preference reversed itself. It is important to note that, in most cases, the lever position preference was not present initially but developed as testing progressed.

Performance at Terminal Delays

The results of the final 20 testing sessions run at the terminal delays are shown in Figure 3. As is clear from the figure, the difference in response accuracy on the DOR and predelay trials which emerged during the previous phase of the experiment maintained itself throughout the 20 sessions. Evaluated by an analysis of variance which treated subjects and blocks as fixed effects, this difference is highly significant [$F(1/90) = 497.6$, $p < .001$]; it is also quite substantial (98.5% vs 63.4% correct responses). None of the animals managed more than 65% responses on DOR trials.

The three animals that participated in this phase of the experiment had previously developed a preference for the left lever, which was assigned to deliver DOR trials during the 20 sessions. Although there was a substantial decline in the percentage choice of DOR trials over the first three blocks of sessions, the animals' preference for the left lever did not reverse itself and indeed showed some recovery during the final block of four sessions (Figure 3). Nor did any of the animals examined individually show a reversal of preference for the left lever.

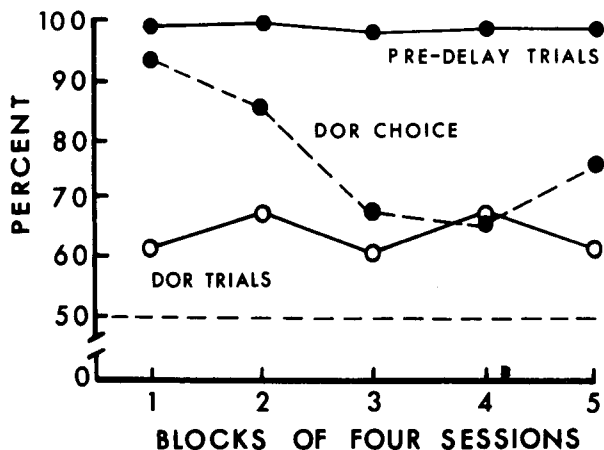


Figure 3. The visual discrimination and leverpress performance over the final 20 sessions, during which the lever that delivered DOR-type trials was maintained on the left for all three animals. The duration of the delay was 32 sec for two of the animals and 64 sec for the third.

Analysis of Response Latencies

Further information regarding the animals' leverpress and discriminative behavior during the last 20 sessions was obtained by analyzing response latencies, measured to the nearest 20 msec. Several repeated measures analyses of variance were performed based on the mean log latencies obtained in each condition. The main effects in these analyses were blocks of sessions, lever position (left or right), trial condition (free or forced), and subjects. The major findings were: (a) The animals responded faster on the left lever than on the right [$F(1/8) = 40.21$, $p < .01$], which probably reflects the animals' position bias rather than the fact that the left lever produced DOR trials. (b) Response latency was significantly longer when a single lever was illuminated (i.e., when the lever response was forced) than when both options were available [$F(1/18) = 17.36$, $p < .01$]. (c) The interaction between lever position (left or right) and trial condition (free or forced) was insignificant. This result is revealing because had the animals learned the relationship between the leverpress response and the nature of the ensuing delay and had they been motivated by the difference in error rate to avoid DOR trials, one would expect unusually long latencies to occur when the animals were forced to respond to the left lever and therefore receive a DOR trial. In statistical terms, a significant Lever Position by Trial Condition interaction would have occurred.

A similar analysis was performed on the projector key responses, i.e., responses to the discriminative stimuli. In one analysis, only correct responses were considered, the relevant factors being type of delay (DOR or predelay), blocks of test sessions, and subjects. (Trial condition, free or forced, was entered as a pseudovisible.) Type of delay was highly significant [$F(1/30) = 15.38$, $p < .01$], responses to S+ being much slower on DOR than on predelay trials. However, this variable interacted with subjects, with two of the three subjects showing a difference in the indicated direction and the third (Hubert) producing the opposite relationship. A similar pattern of results was obtained from an analysis of response latencies on DOR trials only, which included correct and incorrect responses as a factor. The same two animals again produced longer latencies on correct responses than on incorrect choices, whereas Hubert responded somewhat faster on incorrect responses. This animal had by far the shortest response latencies, both on the levers and on the projector keys, possibly too short to reveal the effects of variables that expressed themselves in the other animals.

DISCUSSION

The present results show once more the powerful

and lasting deleterious effect that DOR can impose on heavily overlearned discriminative behavior. The design of the present experiment rules out the possibility that the adverse effect of DOR is due to an associative loss; rather, the responsible mechanism seems to be the loss of incentive value which a DOR occasions. That a DOR reduces the incentive value of a reward object is a hypothesis that has a long history, and recently Ainslie (1975) has shown how certain quantitative realizations of this hypothesis can account for commitment behavior and a variety of other phenomena in psychology and economics.

But even granting that a DOR reduces the incentive value of the reward object, our results are surprising in that an error not only led to the omission of reward but also resulted in a substantial time-out, in the case of two of the animals approximately twice the duration of the DOR itself. The three animals that were run for 20 sessions at the terminal DOR averaged approximately 1% errors on pre-delay trials and 36% errors on DOR trials. This enormous increase in error rate suggests that the DOR must have reduced the incentive value of the reward object to an extremely low level. We have suggested that a DOR may reduce the incentive value of a reward object by two mechanisms, by the delay per se and by an aversive component akin to anticipatory frustration (cf. D'Amato & Cox, 1976). Some evidence for the operation of the latter factor is available in the finding that for two of the three animals latencies of correct responses were longer on DOR than on pre-delay trials, which suggests some degree of reluctance to approach S+ on DOR trials. It should be pointed out that this difference was obtained even though on DOR trials S+ and S- appeared as soon as the leverpress response was executed and therefore while the animal was situated directly before the projector array. In contrast, on pre-delay trials the animal could be anywhere in the experimental chamber when the delay period timed out and the discriminative stimuli appeared.

Despite the marked difference in error rates on pre-delay and DOR trials, none of the four animals learned to avoid pressing the lever that produced DOR trials. Although this result may seem odd at first glance, it is entirely in keeping with the incentive loss hypothesis. Inasmuch as the incentive value of the reward object was so reduced by the DOR that it was incapable of maintaining an overlearned discrimination at a high level, it is not likely to have been sufficient to support the learning of a new discrimination (the leverpress response). Conversely, if there had been sufficient incentive value in the reward that occurred at the end of a trial to reinforce the leverpress response, it should have been sufficient to maintain the visual discriminative response on DOR trials. This line of reasoning suggests the paradox that the greater the error rate on DOR

trials, the less likely it is that the animal will learn to avoid pressing the lever that produces these trials.

The present results raise a caveat to anyone who attempts to produce long-delay discrimination learning in animals. Unless one first demonstrates that the acquisition procedures will support performance of a well-learned discrimination at the value of DOR under examination, the failure to acquire a new discrimination may simply reflect an inadequate incentive value of the reward object rather than an inability on the part of the subjects to form the required association. From the results of the present experiment, it seems unreasonable to expect that under comparable experimental conditions our animals could achieve an asymptotic level of acquisition much higher than 65% if they were exposed to a new discrimination problem at the terminal DOR durations.

It is interesting, in this connection, to compare the present and past performance of the three animals of the present experiment that served in an earlier discrimination acquisition study with DOR (D'Amato & Cox, 1976). Hubert, who in the present experiment was unable to maintain a high performance level on DOR trials with a delay greater than 16 sec, in the earlier study acquired a visual discrimination rapidly with a 1-sec DOR but failed to learn a second visual discrimination when the DOR was increased to 15 sec. Peanuts managed to maintain a high performance level with a 32-sec DOR in the present study. She earlier acquired visual discriminations with DORs of 15 and 30 sec. Finally, Phurp, who reached the longest delay interval in the present experiment, 128 sec, earlier learned a visual discrimination with a 45-sec DOR in only seven sessions of 24 trials each. He also was able to acquire a visual discrimination with a 60-sec DOR, but interestingly enough, after reaching an accuracy level of 90% correct, his performance fell off to only 70%.

The difficulty faced by nonverbal organisms in maintaining high performance levels when the consequences of their behavior are delayed for substantial periods of time is not restricted to DOR. D'Amato and Cox (1976) attempted to maintain discriminative performance by a delayed punishment procedure in which the "punishment" consisted of a lengthy time-out. Despite prolonged training, their animals were unable to maintain a high performance level when the delay reached 1 min or so. Because reward pellets were available only during the delay interval, the latter could not have induced negative emotional responses. They concluded that successful maintenance of instrumental behaviors whose consequences are delayed—which they viewed as a form of foresightful behavior—was probably beyond the reach of most animal species. They pointed out that there is little selective pressure on

animals to develop this type of behavior and that where it does occur in humans it probably is mediated by language or other devices that somehow serve to overcome the loss of incentive value (positive or negative) normally caused by substantial delays.

REFERENCES

- AINSLIE, G. Specious reward: A behavioral theory of impulsiveness and impulse control. *Psychological Bulletin*, 1975, **82**, 463-496.
- BULLOCK, M. A., & RICHARDS, R. W. Disruption of fixed ratio performance as a function of reinforcement delay. *Bulletin of the Psychonomic Society*, 1973, **1**, 49-52.
- D'AMATO, M. R., & COX, J. K. Delay of consequences and short-term memory in monkeys. In D. L. Medin, W. A. Roberts, & R. T. Davis (Eds.), *Processes in animal memory*. Hillsdale, N.J: Erlbaum, 1976.
- DAVENPORT, J. The interaction of magnitude and delay of reinforcement in spatial discrimination. *Journal of Comparative and Physiological Psychology*, 1962, **55**, 267-273.
- FERSTER, C. B., & HAMMER, C. Variables determining the effects of delay of reinforcement. *Journal of the Experimental Analysis of Behavior*, 1965, **8**, 243-254.
- LOGAN, F. A. Decision making by rats: Delay versus amount of reward. *Journal of Comparative and Physiological Psychology*, 1965, **59**, 1-12.
- MISHKIN, M., & WEISKRANTZ, L. Effects of delaying reward on visual-discrimination performance in monkeys with frontal lesions. *Journal of Comparative and Physiological Psychology*, 1958, **51**, 276-281.
- MOWRER, O. H. *Learning theory and behavior*. New York: Wiley, 1960.
- RENNER, K. E. Delay of reinforcement: A historical review. *Psychological Bulletin*, 1964, **61**, 341-361.

(Received for publication July 8, 1976;
revision accepted October 28, 1976.)